IN THE CLAIMS

Please amend the claims as follows:

1. (Currently Amended) A document classification system for classifying a document based on contents of the document of which contents contains a plurality of items, said document classification system comprising:

inputting means for inputting document data corresponding to the document;

displaying means for displaying the items of the document data input by said inputting means;

designating means for designating a ~~portion~~ plurality of the items displayed in said displaying means;

converting means for converting the document data into converted data so that the converted data contains only data corresponding to the items designated by said designating means which are a portion less than all of the items of the document data; and

classifying means for classifying the document by using the converted data produced by said converting means.

2. (Original) The document classification system as claimed in 1, wherein said classifying means includes document vector producing means for producing a feature vector representing a feature of the converted data so as to classify the document in accordance with the feature vector produced by said document vector producing means.

3. (Original) The document classification system as claimed in 1, wherein said converting means includes separation sign inserting means for inserting a predetermined sign between sets of data corresponding to the items so as to facilitate separation of each data corresponding to each item in the converted data.

4-6 (Canceled).

7. (Currently Amended) A processor readable medium storing program code causing a computer to classify a document based on contents of the document of which contents contains a plurality of items, comprising:

first program code means for inputting document data corresponding to the document;

second program code means for displaying the items of the document data input by said first program code means;

third program code means for designating a ~~portion~~ plurality of the items displayed by the second program code means;

fourth program code means for converting the document data into converted data so that the converted data contains only data corresponding to the items designated by the third program code means which are a portion less than all of the items of the document data; and

fifth program code means for classifying the document by using the converted data produced by the fourth program code means.

8. (Previously Presented) The processor readable medium as claimed in 7, wherein the fifth program code means includes sixth program code means for producing a feature vector representing a feature of the converted data so as to classify the document in accordance with the feature vector.

9. (Previously Presented) The processor readable medium as claimed in 7, wherein the fourth program code means includes seventh program code means for inserting a predetermined sign between sets of data corresponding to the items so as to facilitate separation of each data corresponding to each item in the converted data.

10. (Currently Amended) A document classification system for classifying a document according to contents of the document, said document classification system comprising:

input means for inputting document data of the document;

analyzing means for analyzing the document data so as to obtain analysis information;

vector producing means for producing a document feature vector with respect to the document data based on the analysis information;

transforming function calculating means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected with a dimensional number different from a dimensional number of the document feature vector, the transforming function calculating means calculating the representation transforming function by using an inner product calculated between the document feature vectors;

vector transforming means for transforming the document feature vector by using the representation transforming function;

classification means for classifying the document based on similarity between the document feature vectors transformed by the vector transforming means; and

classification result storing means for storing a result of classification performed by the classification means.

11. (Original) The document classification system as claimed in 10, further comprising inner product calculating means for calculating an inner product between the document feature vectors, wherein said representation transforming function calculating means calculates the representation transforming function by using the inner product.

12. (Previously Presented) A document classification system for classifying a document according to contents of the document, said document classification system comprising:

input means for inputting document data of the document;

analyzing means for analyzing the document data so as to obtain analysis information;

vector producing means for producing a document feature vector with respect to the document data based on the analysis information;

transforming function calculating means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

vector transforming means for transforming the document feature vector by using the representation transforming function;

classification means for classifying the document based on similarity between the document feature vectors transformed by the vector transforming means;

classification result storing means for storing a result of classification performed by the classification means;

inner product calculating means for calculating an inner product between the document feature vectors, wherein said representation transforming function calculating means calculates the representation transforming function by using the inner product; and

document similarity information setting means for setting document similarity setting information including data representing an author of the document and a date of production of the document, wherein said representation transforming function calculating means calculates the representation transforming function by using the inner product and the document similarity information.

13. (Original) The document classification system as claimed in 10, further comprising:

vector storing means for storing the document feature vector produced by said vector producing means; and

transforming function storing means for storing the representation transforming function calculated by said representation transforming function calculating means.

14. (Previously Presented) A document classification system for classifying a document according to contents of the document, said document classification system comprising:

input means for inputting document data of the document;

analyzing means for analyzing the document data so as to obtain analysis information;

vector producing means for producing a document feature vector with respect to the document data based on the analysis information;

transforming function calculating means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

vector transforming means for transforming the document feature vector by using the representation transforming function;

classification means for classifying the document based on similarity between the document feature vectors transformed by the vector transforming means;

classification result storing means for storing a result of classification performed by the classification means; and

vector correcting means for correcting the document feature vector before the

document feature vector is transformed by said vector transforming means, a correction being

performed by processing one of the document feature vector and a feature dimension

constituting the document feature vector in accordance with a rule established by

characteristics of words extracted by said analyzing means.

15. (Original) The document classification system as claimed in 14, further

comprising transforming function correcting means for correcting the representation

transforming function calculated by said transforming function calculating means when the

feature dimension is changed due to a correction of the document feature vector by said

vector correcting means so that the document feature vector is transformed by said vector

transforming means in accordance with the changed feature dimension.

16. (Previously Presented) A document classification system for classifying a

document according to contents of the document, said document classification system

comprising:

input means for inputting document data of the document;

analyzing means for analyzing the document data so as to obtain analysis information;

vector producing means for producing a document feature vector with respect to the

document data based on the analysis information;

transforming function calculating means for calculating a representation transforming

function used for projecting the document feature vector onto a space in which similarity

between the document feature vectors is reflected;

vector transforming means for transforming the document feature vector by using the

representation transforming function;

classification means for classifying the document based on similarity between the document feature vectors transformed by the vector transforming means;

classification-result storing means for storing a result of classification performed by the classification means;

transforming function correction instructing means for sending an instruction regarding a process to be applied on a feature dimension of the representation transforming function; and

transforming function correcting means for correcting the representation transforming function based on a content of the instruction sent from said transforming function correction instructing means.

17. (Original) The document classification system as claimed in 16, wherein the process indicated in the content of the instruction is performed by using data of an arbitrary document vector.

18. (Original) The document classification system as claimed in 16, wherein the process indicated in the content of the instruction is performed by using the document feature vectors.

19. (Original) The document classification system as claimed in 16, wherein the process indicated in the content of the instruction is performed by using the analysis information obtained by said analyzing means.

20. (Original) The document classification system as claimed in 16, wherein the process indicated in the content of the instruction is performed by using the result of classification stored in said classification-result storing means.

21. (Previously Presented) A document classification system for classifying a document according to contents of the document, said document classification system comprising:

input means for inputting document data of the document;

analyzing means for analyzing the document data so as to obtain analysis information;

vector producing means for producing a document feature vector with respect to the document data based on the analysis information;

transforming function calculating means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

vector transforming means for transforming the document feature vector by using the representation transforming function;

classification means for classifying the document based on similarity between the document feature vectors transformed by the vector transforming means;

classification result storing means for storing a result of classification performed by the classification means;

an initial cluster centroid designating means for designating an initial cluster centroid; and

initial cluster centroid registering means for registering the initial cluster centroid designated by said initial cluster centroid designating means,

wherein said classification means classifies the document in accordance with the initial cluster centroid registered by said initial cluster centroid registering means.

22. (Original) The document classification system as claimed in 21, wherein the initial cluster centroid designated by said initial cluster centroid designating means is arbitrary document vector data.

23. (Original) The document classification system as claimed in 21, wherein the initial cluster centroid designated by said initial cluster centroid designating means is the document feature vector.

24. (Original) The document classification system as claimed in 21, wherein the initial cluster centroid designated by said initial cluster centroid designating means is the analysis information obtained by said analyzing means.

25. (Original) The document classification system as claimed in 21, wherein the initial cluster centroid designated by said initial cluster centroid designating means is the result of classification stored by said classification-result storing means.

26-41 (Canceled).

42. (Currently Amended) A processor readable medium storing program code causing a computer to classify a document according to contents of the document, comprising:

first program code means for inputting document data of the document;

second program code means for analyzing the document data so as to obtain analysis information;

third program code means for producing a document feature vector with respect to the document data based on the analysis information;

fourth program code means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected with a dimensional number different from a dimensional number of the document feature vector, the fourth program code means calculating the representation transforming function by using an inner product calculated between the document feature vectors;

fifth program code means for transforming the document feature vector by using the representation transforming function;

sixth program code means for classifying the document based on similarity between the document feature vectors transformed by the fifth program code means; and

seventh program code means for storing a result of classification performed by the classification means.

43. (Original) The processor readable medium as claimed in 42, further comprising eighth program code means for calculating an inner product between the document feature vectors, wherein the representation transforming function is calculated by using the inner product.

44. (Previously Presented) A processor readable medium storing program code causing a computer to classify a document according to contents of the document, comprising:

first program code means for inputting document data of the document;

second program code means for analyzing the document data so as to obtain analysis information;

third program code means for producing a document feature vector with respect to the document data based on the analysis information;

fourth program code means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

fifth program code means for transforming the document feature vector by using the representation transforming function;

sixth program code means for classifying the document based on similarity between the document feature vectors transformed by the fifth program code means;

seventh program code means for storing a result of classification performed by the classification means;

eighth program code means for calculating an inner product between the document feature vectors, wherein the representation transforming function is calculated by using the inner product; and

ninth program code means for setting document similarity setting information including data representing an author of the document and a date of production of the document, wherein the representation transforming function is calculated by using the inner product and the document similarity information.

45. (Original) The processor readable medium as claimed in 42, further comprising:

tenth program code means for storing the document feature vector produced by the third program code means; and

eleventh program code means for storing the representation transforming function

calculated by the fourth program code means.

46. (Previously Presented) A processor readable medium storing program code

causing a computer to classify a document according to contents of the document,

comprising:

first program code means for inputting document data of the document;

second program code means for analyzing the document data so as to obtain analysis

information;

third program code means for producing a document feature vector with respect to the

document data based on the analysis information;

fourth program code means for calculating a representation transforming function

used for projecting the document feature vector onto a space in which similarity between the

document feature vectors is reflected;

fifth program code means for transforming the document feature vector by using the

representation transforming function;

sixth program code means for classifying the document based on similarity between

the document feature vectors transformed by the fifth program code means;

seventh program code means for storing a result of classification performed by the

classification means; and

eighth program code means for correcting the document feature vector before the

document feature vector is transformed by the fifth program code means, a correction being

performed by processing one of the document feature vector and a feature dimension

constituting the document feature vector in accordance with a rule established by

characteristics of words extracted by the second program code means.

13

47. (Previously Presented) The processor readable medium as claimed in 46, further comprising ninth program code means for correcting the representation transforming function calculated by the fourth program code means when the feature dimension is changed due to a correction of the document feature vector by the eighth program code means so that the document feature vector is transformed by the fifth program code means in accordance with the changed feature dimension.

48. (Previously Presented) A processor readable medium storing program code causing a computer to classify a document according to contents of the document, comprising:

first program code means for inputting document data of the document;

second program code means for analyzing the document data so as to obtain analysis information;

third program code means for producing a document feature vector with respect to the document data based on the analysis information;

fourth program code means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

fifth program code means for transforming the document feature vector by using the representation transforming function;

sixth program code means for classifying the document based on similarity between the document feature vectors transformed by the fifth program code means;

seventh program code means for storing a result of classification performed by the classification means;

eighth program code means for sending an instruction regarding a process to be applied on a feature dimension of the representation transforming function; and

ninth program code means for correcting the representation transforming function based on a content of the instruction sent by the eighth program code means.

49. (Previously Presented) A processor readable medium storing program code causing a computer to classify a document according to contents of the document, comprising:

first program code means for inputting document data of the document;

second program code means for analyzing the document data so as to obtain analysis information;

third program code means for producing a document feature vector with respect to the document data based on the analysis information;

fourth program code means for calculating a representation transforming function used for projecting the document feature vector onto a space in which similarity between the document feature vectors is reflected;

fifth program code means for transforming the document feature vector by using the representation transforming function;

sixth program code means for classifying the document based on similarity between the document feature vectors transformed by the fifth program code means;

seventh program code means for storing a result of classification performed by the classification means;

eighth program code means for designating an initial cluster centroid; and

ninth program code means for registering the initial cluster centroid designated by the eighth program code means,

wherein the document is classified in accordance with the initial cluster centroid

registered by the ninth program code means.

50. (Currently Amended) A document classification system for classifying a

document based on contents of the document of which contents contains a plurality of items

irrespective of chapters, clauses, sentences and paragraphs of the document, said document

classification system comprising:

inputting means for inputting document data corresponding to the document;

displaying means for displaying the items of the document data input by said

inputting means;

designating means for designating a ~~portion~~ plurality of the items displayed by said

displaying means;

converting means for converting the document data into converted data so that the

converted data contains only data corresponding to the items designated by said designating

means which are a portion less than all of the items of the document data; and

classifying means for classifying the document by using the converted data produced

by said converting means.